ATLAS Distributed Computing Operations Shift Team Experience

Kaushik De[1], Xavier Espinal[2,9], Alessandra Forti[3], Elena Korolkova[4], Kai Leffhalm[5], Peter Love[6], Jaroslava Schovancová[7], Yuri Smirnov[8] for the ATLAS Collaboration

[1] Department of Physics, The University of Texas at Arlington, Arlington TX, United States of America
[2] Port d'Informació Científica (PIC), Universitat Autònoma de Barcelona, Edifici D, ES-08193 Bellaterra (Barcelona), Spain
[3] School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
[4] Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom
[5] Deutsches Elektronen-Synchrotron, Zeuthen, Germany
[6] Physics Department, Lancaster University, Lancaster, United Kingdom
[7] Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, CZ-18221 Prague 8, Czech Republic
[8] Physics Department, Brookhaven National Laboratory, Upton NY, United States of America

E-mail: `schovan@fzu.cz`

Abstract. Description of ATLAS Distributed Computing Operations Shift (ADCoS) duties and available monitoring tools along with the operational experience is presented in this paper.

1. Introduction

ATLAS Distributed Computing Operations Shift (ADCoS) Team was organized in 2008 to support 24/7 simulation production of the ATLAS experiment [1], data reprocessing and data management operations at hundreds of computing centers located world-wide. ATLAS uses PanDA (Production and Distributed Analysis) [2] and DQ2 (Distributed Data Management) [3] system for distributed backend workflow, along with various monitoring tools (for example, ARDA-dashboard, Pandamon) and work injection systems (ProdSys) to manage automated data distribution and processing using the ATLAS framework Athena [4]. PanDA continuously runs 200k-300k jobs per day all over the world on the Grid, and DQ2 has reached more than 9 GB/s integrated transfer rate over all ATLAS sites. To provide reliable and efficient quality of service with such a complex system, we need a 24/7 testing and monitoring team. We have set up and run such a team, ADCoS, for the past two years. The ADCoS team is distributed world-wide and works remotely from their home institutions, and remotely from the software developers. In this paper, we describe the structure of ADCoS, the lessons learned, and the numerous tools used to work efficiently.

[9] Also at Institut de Física dAltes Energies, IFAE, Edifici Cn, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain

2. Organisation

ATLAS Distributed Computing Operations Shift (ADCoS) is a world-wide distributed Team organised in three different timezones: the Asia-Pacific timezone (00-08 hrs CET), the Europe timezone (08-16 hrs CET), and the US timezone (16-24 hrs CET). Every timezone has a Shift Captain, who acts as a contact person for shift training and coordinates timezone shift efforts.

ADCoS Team monitors ATLAS Distributed Computing (ADC) activities every single day in a year. There are three flavours of shifters: Trainees, Seniors, and Experts. Every single day there are three Seniors on duty, one per each timezone. Every day except Sunday there are three shift slots allocated for Trainees. There is one Expert shift slot per day.

3. Training

ADCoS Training consists of two components: the Tutorial session, and the Trainee Shift. There are two Tutorials organised every year. The tutorial is a half-day-long session that covers various crucial activities of the ADCoS Team. The aim of the Tutorial is to give a brief overview of ADCoS duties. Tutorial usually takes place as a part of ATLAS Software & Computing Workshop at CERN. The tutorial session is broadcasted as a video-conference. Since recently, the tutorial session is videorecorded and available in archive of CERN Audiovisual service.

Once the Trainee shifter attends the tutorial session he/she can sign up for Trainee Shift. Trainee Shifter fulfills common ADCoS tasks and communicates with Senior and Expert Shifters, which helps to effectively pass knowledge among the ADCoS Team members.

4. Communication

ADCoS Team members communicate mainly through mailing list, and through ADC Virtual Control Room. The ADC Virtual Control Room needs to accomodate up to a hundred ADCoS Team members during a certain period of time, there are about 10 ADCers connected to ADC Virtual Control Room at any time. ADC Virtual Control Room is used by three shift teams at the same time, instant message delivery is of crucial importance for us. Communication history is logged and recorded, so that very recent history is available to shifters on duty, and multiple conversation threads are available. At the same time, the instant communication cannot affect security of any of the ATLAS institutions.

Variety of tools and technologies have been tried, starting with Skype group chat room, through EVO and Google wave, we end up with Jabber group chat room. On the way we suffered from scaling issues resulting in message delivery delayed by tens of minutes to tens of hours, through clients with enormous memory consumption, to lack of history logging.

5. ATLAS Distributed Computing monitoring

ATLAS Distributed Computing is a very complex system. To provide reliable and efficient quality of service with such a complex system, we need a 24 hours per day 7 days per week testing and monitoring team. Issue spotted on the monitoring pages by the ADCoS Shifter on duty is investigated. Instructions for ADCoS Shifters are summarised on a TWiki page. ADCoS Shifter on duty follows the procedure and collects as much information as possible about this issue. When all necessary information is gathered from monitoring tools and available logs, he reports that issue in a ticket. In the meantime, failing service is excluded from production. ADCoS monitor two areas of activities: Distributed Data Management, and Production jobs. Under special circumstances, ADCoS monitor also ADC Central services.

Distributed Data Management is monitored on the DDM Dashboard [5]. The software it provides is able to collect information regarding data transfers from different sources and locations, and exposes to the grid end user a processed view of this data. It provides essential information about both successful and failed data transfers. This information can be passed to administrator of any failing DDM-related service.

ADC production activites consist of data processing at Tier0, data reprocessing at Tier1s, MC simulations at Tier1s and Tier2s, and Physics Group production at Tier1s and Tier2s. ADCoS monitors all these activities but data processing at Tier0. Production jobs monitoring consists of two tools: Panda monitor [6] and ProdSys dashboard. These monitoring tools provide complementary information about production tasks and jobs, e.g. job status, job logs, task efficiency, etc. ADCoS Shifters can easily spot failing jobs and dig deeper into job logs to reveal the cause of such a failure. Once the cause of failure is unveiled, issue is either reported in a bug report, or escalated to Team with more expertise in the particular field.

ADC Central services are computing services and resources based at CERN of crucial importance to ATLAS Distributed Computing, e.g. databases, Castor storage, and various ADC products supporting DDM and Production operations. ADC Central services are monitored by a different ADC Shift Team, the Comp@P1 Shift team, by default. Under special circumstances, when there is no Comp@P1 Shifter on duty, e.g. during CERN closure period, ADCoS monitor ADC Central services in addition to their duties. ADC Central services status is shown on SLS (Service Level Status) monitors. When SLS monitors show degraded service, issue is escalated to ADC Manager on Duty.

The long-term goal of the ADC Operations is to minimize necessary manual intervention of Operations Shift Teams, therefore some of the subsystem failures are automatically reported to service providers, and the failing service is temporarily taken out of production. Once the issue is fixed, ADCoS Shifters perform particular service test.

6. Issue Reports and Follow-ups

Every failure of any subsystem of the three areas of activities is reported so that proper action is taken by the service provider and the issue is fixed. Software bugs are reported to responsibles via LCG Savannah portal. Global Grid User Support (GGUS) system is used to report site-related issues to the sites, and as a communication channel between ADC Operations and sites.

Anytime ADCoS Shifter takes an action, this action is to be recorded as an eLog entry. The ADC eLogbook contains history of any kind of action taken by ADC Operations members. At the end of the daily shift ADCoS Shifter submits a daily report with the most important issues of the day so that the following ADCoS Shifter on duty is informed and can continue in problem chasing from the point where previous ADCoS Shifter stopped. The daily reports are aggregated in weekly reports presented at ADCoS Weekly meetings.

7. Conclusions

The ATLAS Distributed Computing Operations Shift Team ensures reliability and quality of service of the ATLAS Distributed Computing system for almost 3 years. ADCoS Team monitors ADC subsystems 24 hrs per day, 7 days per week. ADCoS Team reports or escalates observed issues, and perform service tests when necessary.

8. Acknowledgements

The Port dInformació Científica (PIC) is maintained through a collaboration between the Generalitat de Catalunya, CIEMAT, IFAE and the Universitat Autònoma de Barcelona. This work was supported in part by grant FPA2007-66152-C02-00 from the Ministerio de Educación y Ciencia, Spain. Jaroslava Schovancová greatfully appreciates support from the Academy of Sciences of the Czech Republic. Support from the grant LA08032 of the MEYS (MSMT), Czech Republic, and from ACEOLE: A Marie Curie ITN Project is greatly acknowledged. The work was supported by the grant SVV-2010-261 309 of the Charles University in Prague.

References

[1] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” JINST **3** (2008) S08003.

[2] T. Maeno, “PanDA: Distributed production and distributed analysis system for ATLAS,” J. Phys. Conf. Ser. **119** (2008) 062036.

[3] M. Branco *et al.*, “Managing ATLAS data on a petabyte-scale with DQ2,” J. Phys. Conf. Ser. **119** (2008) 062017.

[4] P. Calafiura *et al.*, “The Athena control framework in production, new developments and lessons learned,” in Proceedings of the CHEP’04 Conference, Interlaken, Switzerland, 27 Sept. - 1 Oct. 2004, pp. 456-458

[5] J. Andreeva *et al.*, “Dashboard applications to monitor experiment activities at sites,” J. Phys. Conf. Ser. **219** (2010) 062003.

[6] Klimentov, A., Nevski, P., Potekhin, M., Wenaus, T., “The ATLAS PanDA Monitoring System and its Evolution”, to appear in Proceedings of the CHEP 2010 International Conference